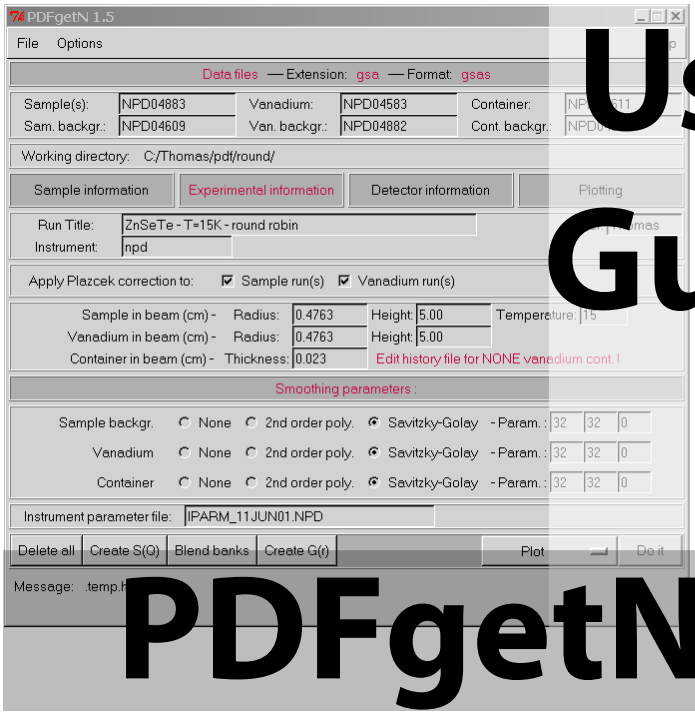

Users Guide

PDFgetN 1.6

written by

Thomas Proffen
Email: tproffen@lanl.gov

Peter Peterson
Email: petersonpf@ornl.gov

Matthias Gutmann
Email: M.J.Gutmann@rl.ac.uk

Simon Billinge
Email: billinge@pa.msu.edu

Document created: December 14, 2005
Current version: **1.6.2**

Preface

Disclaimer

By downloading the program *PDFgetN*, you agree to the terms and conditions concerning its use specified in the license agreement that is provided as part of the distribution. End users wishing to make commercial use of the software must contact Libraries, Computing & Technology, Michigan State University (MSU), East Lansing, MI 48824; (517)353-0722 prior to any commercial distribution to discuss terms. The Software is provided to End User by MSU on an as is basis. No user support is provided or implied. MSU makes no warranty, express or implied to end user or to any other person or entity. Specifically, MSU makes no warranty of merchantability or fitness for a particular purpose of the software. MSU will not be liable for special, incidental, consequential, indirect or other similar damages, even if MSU or its employees have been advised of the possibility of such damages, regardless of the form of the claim.

Acknowledgments

Financial support from the National Science Foundation through the grants DMR-9700966 and CHE-9633798 as well as the Center for Fundamental Materials Research (CFMR) is gratefully acknowledged. MG acknowledges financial support from the Swiss National Science Foundation. This work has benefited from the use of the Intense Pulsed Neutron Source at Argonne National Laboratory funded by the U.S. Department of Energy, BES-Materials Science, under Contract W-31-109-Eng-38 and from the use of the Lujan Center at Los Alamos Neutron Science Center, funded by Department of Energy Office of Basic Energy Sciences and Los Alamos National Laboratory funded by Department of Energy under contract W-7405-ENG-36. We thank W. Dmowski for further suggestions and his extensive testing of *PDFgetN*.

Using PDFgetN

Publication of results totally or partially obtained using the program *PDFgetN* should state that *PDFgetN* was used and contain the following reference:

> PETERSON, P.F., GUTMANN, M., PROFFEN, TH. & BILLINGE, S.J.L. (2000) "PDFgetN: A User-Friendly Program to Extract the Total Scattering Structure Function and the Pair Distribution Function from Neutron Powder Diffraction Data" *J. Appl. Cryst.*, **33**, 1192

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 What is PDFgetN ?

Now you might have installed *PDFgetN* and can start it simply by typing `PDFgetN` if you are running UNIX, or double clicking on the *PDFgetN* icon in case you work with Windows. However, in case you are not a PDF wizard here is a little introduction into the PDF method first. The remaining users guide explains the various features of *PDFgetN* more detail.

The determination of crystal structures is an important part of chemistry, physics and of course crystallography. Conventional structure determination is based on the analysis of the intensities and positions of Bragg reflections which only allows one to determine the long range *average* structure of the crystal. Only one-body information such as atomic positions, site occupancies and temperature factors can be extracted. Determination of the *average* structure based on powder diffraction data is now routinely done using the Rietveld (Rietveld, 1969) method which is very similar to the full profile refinement of the atomic pair distribution function (PDF) as we will see later. It should be kept in mind that the analysis of Bragg scattering assumes a prefect long range periodicity of the crystal. However, many materials are quite disordered and even more important the key to the deeper understanding of their properties is the study of deviations from the *average* structure or the study of the *local* atomic arrangements. Deviations from the *average* structure result in the occurrence of diffuse scattering which contains information about two-body interactions (Welberry and Butler, 1995; Frey, 1995). In recent years the analysis of diffuse scattering from single crystals as well as powders using computer simulations have made great advances, in particular using the Monte Carlo (MC) and the Reverse Monte Carlo (RMC) technique (Welberry and Butler, 1995; Nield et al., 1995; Welberry and Proffen, 1998; Proffen and Welberry, 1998, 1997).

Another method to reveal the *local* structure of a crystals is the analysis of the PDF. This method is long known in the field of studying short range order in liquids and glasses but has recently been applied to crystalline materials (Egami, 1998; Billinge and Egami, 1993). The PDF is obtained from the powder diffraction data via a simple Fourier transform of the normalized scattering intensity $S(Q)$:

$$G(r) = 4\pi r[\rho(r) - \rho_0] = \frac{2}{\pi}\int_0^\infty Q[S(Q) - 1]\sin(Qr)dQ, \tag{1.1}$$

where $\rho(r)$ is the microscopic pair density, ρ_0 is the average number density and Q is the magnitude of the scattering vector. For elastic scattering it is $Q = 4\pi \sin(\theta)/\lambda$, with 2θ being the scattering angle and λ the wavelength of the radiation used. Since the PDF contains Bragg and diffuse scattering, the information about *local* arrangements is preserved. The PDF can be understood as a bond-length distribution between all pairs of atoms i and j within the crystal (up to a maximum distance), however each contribution has a weight corresponding to the scattering power of the two atoms involved. In order to carry out the Fourier transform in (1.1) we would need to measure data up to $Q = \infty$, which of course is not possible. Thus the termination at a value of $Q = Q_{max}$ will cause so-called *termination ripples* in the PDF which, however, can easily be modelled. With the availability of modern synchrotron and neutron sources it is possible to collect powder diffraction data up to high values in Q.

Once, a PDF is obtained, there are several ways to model it. Most desirable is a method based on a structural model. Two programs that work directly with the output of *PDFgetN* will be mentioned here: (1) PDFFIT, a real-space Rietveld program (Proffen and Billinge, 1999) which is used for refining a small structural model and (2) DISCUS (Proffen and Neder, 1997) using a Reverse Monte Carlo approach (McGreevy and Pusztai, 1988) e.g. to yield structural parameters from a large model "box". The modelling of the total scattering function $S(Q)$ also becomes increasingly important. This is partly because the data is corrected for background and multiple scattering effects and thus the background becomes better defined in Q regions where Bragg peaks overlap. This of course could only be a very short overview of the PDF technique, much more information can be found in a recent review article by Proffen et al. (2003) or an upcoming book about the PDF technique by Egami and Billinge (2003).

After talking about the PDF and its modelling very generally, the question remains what does *PDFgetN* actually do ? In a first step, the measure data are corrected for background scattering, container scattering as well as the incident spectrum according to Eq. 1.2.

$$I = \frac{(S - S_B) - \alpha(C - C_B)}{V - V_B} \tag{1.2}$$

Here S is the sample scattering, C is the scattering from an empty sample container and V is the scattering from a vanadium sample which characterizes the incident spectrum. S_B, C_B and V_B are the background spectra for the corresponding runs. If all measurements were made using the same ancillary equipment, all the background runs refer to a single measurement. The factor α takes into account the absorption of the sample. Some more discussion about characterization runs is given in section 3.1. Next, the combined data I need to be normalized and corrected to obtain $S(Q)$. This involves the interaction with several programs that calculate the absorption, multiple scattering and Plazcek (inelasticity) correction, and combine the runs to yield a PDF which forms the basis for a detailed analysis of the local structure. *PDFgetN* allows users to obtain $S(Q)$ and PDF from time-of-flight neutron powder diffraction data from a wide variety of neutron powder diffractometers. It is easy to use with a graphical user interface and embedded data such as neutron scattering lengths. The data analysis programs have been derived from the GLASS package (Price, undated). A key feature of *PDFgetN* is that it saves data analysis parameters for a particular $S(Q)$ and PDF, the so-called data-analysis history, that can be reloaded into *PDFgetN*. The history contains the file names of the

underlying data files and the parameters used to obtain the $S(Q)$ and PDF. Another feature is the possibility to visualize the result of each analysis step with a variety of options. The program also calculates different quality factors (Peterson et al., 2003b) for the resulting $S(Q)$ as well as the PDF, $G(r)$.

Finally the question arises which diffractometers are supported by *PDFgetN* ? The simple answer is, any neutron time-of-flight powder diffractometer that can produce GSAS file is compatible with *PDFgetN*. Some native formats for selected instruments are also supported, for details see section A.1.

1.2 What is new ?

If you have used *PDFgetN* before, you might ask what is new in this version of the program. Below you find sections describing the major changes in the different releases of *PDFgetN*.

Version 1.6

The most notable new feature is the use of a new font, giving a nicer appearance. In addition to a number of bug fixes, the following new features have been added:

- The processed $S(Q)$ can now be saved as GSAS file, so it is possible to use $S(Q)$ for Rietveld refinements. Select *Save GSAS* under the *File* menu.
- The button *Automatic normalization* next to the effective sample density will now determine the correct density to obtain $S(Q) = 1$ at high Q. More details can be found in section 2.4.
- *PDFgetN* now has a feature that allows one to apply an arbitrary correction to $S(Q)$ (see 3.4).
- In addition to the SEPD example files, a new tutorial processing NPDF data files has been included.

Version 1.5

This is the first revision of the users guide since version 1.3 of the program, so all changes referred to in this section are changes between versions 1.3 and 1.5. So the users guide you are currently reading is a new feature of this version of the program. Another difference one will notice is the new self-extracting installer for Windows. If installing *PDFgetN* under Windows, no other installation of supporting software is required. Details about the installation procedure are discussed in appendix B.1.

PDFgetN 1.5 now fully supports the input of GSAS type data files, making this program compatible with almost any neutron time-of-flight powder diffractometer in the world. Even though the previous native file formats for some selected instruments are still available, we strongly recommend using GSAS as the data file format. The plotting features have now been integrated in the main *PDFgetN* panel and in addition to all intermediate files, one can now also plot the values of the different corrections applied (see 2.7). Data files can now be selected from a list by *right-clicking* into the corresponding data files (see 2.2). The Plazcek correction for the sample and/or vanadium run can now be disabled from the GUI (see 2.4) and the final grid

size for $S(Q)$ can now be entered from the GUI as well (see 2.6). In addition the $S(Q)$ values can now be matched to a reference bank before the blending step. Finally two new features allow the calculation of quality factors as well as an automatic optimization of the resulting PDF, $G(r)$ (see 3.3 and Peterson et al. (2003b)).

1.3 More information and other programs

To find out about recent updates of *PDFgetN* or to get further information visit the *PDFgetN* WWW homepage at the following site:

http://www.totalscattering.org/programs/PDFgetN/

This manual will also mention the refinement and modelling programs *PDFFIT* and *DISCUS* as well as the plotting program *KUPLOT*. All of these programs are designed to work with *PDFgetN* output and are part of the *Diffuse* package of programs. For more information and to download these programs, please visit the *Diffuse* homepage at

http://www.totalscattering.org/programs/discus/ or
http://www.uni-wuerzburg.de/mineralogie/crystal/discus/

If all else fails and you are stuck with a problem that seems unsolvable, please contact one of the authors who are more than happy to try and help solve your problem. Also any comments or wishes for future versions are more than welcome.

Chapter 2

Using PDFgetN

In this chapter we will discuss in detail how to use *PDFgetN*. The program window consists of three sections: The *file input* fields on the top. The middle section contains four different information panels: *sample information*, *experimental information*, *detector bank information* and *plotting controls*. The different panels can be activated by clicking on the corresponding tab in the main window of *PDFgetN*. The bottom part contains buttons and controls to make the program actually do something. In the following sections, each section and panel of the program will be discussed in detail.

2.1 Starting PDFgetN

First you need to start the program. In case of Unix/Linux, simply enter `PDFgetN` and provided the installation location is in your path, the program will start. Alternatively you can give the complete path, e.g.

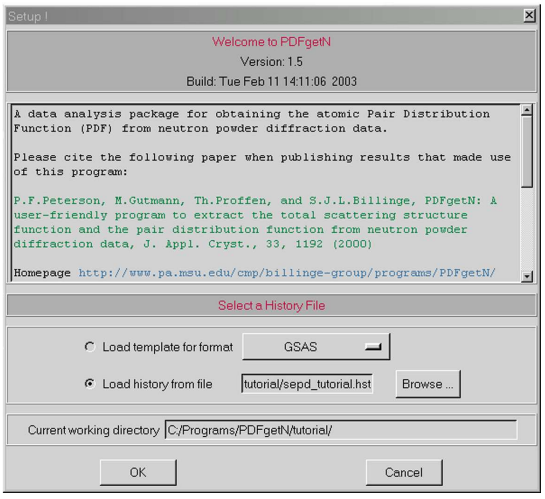

Figure 2.1: Startup screen of *PDFgetN*.

`/usr/special/bin/PDFgetN-1.6/PDFgetN.pl`. If you are running Windows and have installed the desktop icons, simply double click on the *PDFgetN* icon. If no desktop icon is installed, browse to the "binary" directory of the *PDFgetN* installation directory and double click `PDFgetN.exe`. In either case, the startup screen of *PDFgetN* will appear as shown in Figure 2.1. At this point you can either start from scratch by selecting a template from the drop down menu. The template to select is given by the format of the powder diffraction data that will be used. Details about the supported data file formats are given in Appendix A.1. In case a template is used, the desired working directory also needs to be specified in the bottom input field. Note that even under Windows, the directory names are separated by / as common under Unix. The second option is to load a previously created file by either typing in the name or using the *browse* button. Make sure the button next to the desired selection is highlighted. In our example we load the file *sepd_tutorial.hst* which can be found in the *tutorial* directory of the program directory. If you are working with a shared installation, you might not have write permission in the tutorial directory. In this case, simply copy all the files from the tutorial directory to a local directory. After we click *ok*, the tutorial data are read and we are ready to go. Sometimes there might be a third option, *load temporary history file*. In this case the selected working directory contains a history file from a previous *PDFgetN* session. The program is always keeping a temporary history file reflecting the current settings, which allows one to simply continuing at the point *PDFgetN* was terminated last time.

2.2 Data file input

The top section of the *PDFgetN* window allows one to enter the names of the data and characterization run files. This section is highlighted in Figure 2.2. The input in the top row corresponds to the sample S,

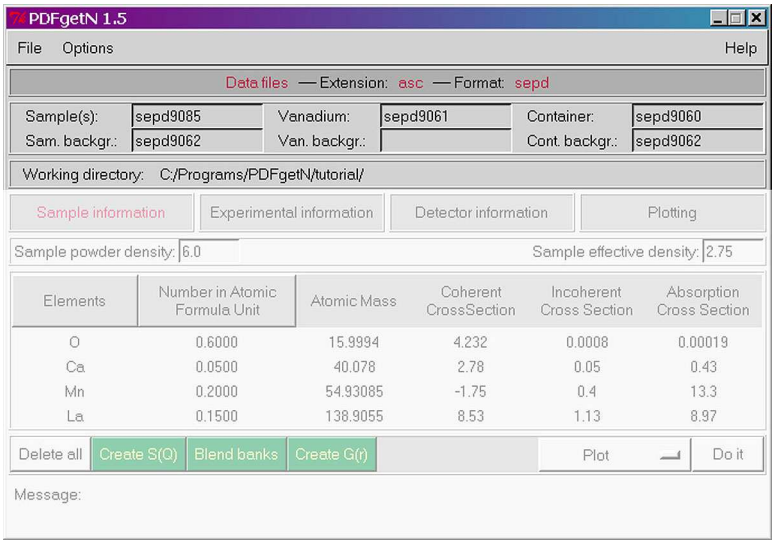

Figure 2.2: File input area of *PDFgetN*.

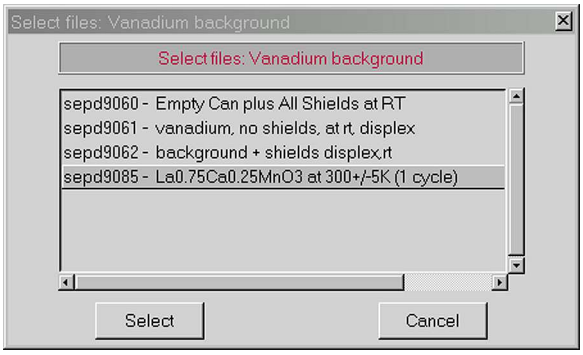

Figure 2.3: File select dialog of *PDFgetN*.

vanadium V and container run C. The row below has input fields for the corresponding backgound runs S_B, V_B and C_B, respectively. The mimimum set of runs are a sample and a vanadium runs, the other fields can be left blank. Some experimental considerations to obtain a high quality PDF are discussed in section 3.1. Sometimes, one wants to merge several runs. This can be done by entering a list of filenames separated by commas. Also note, that the filenames are entered *without* file extension. All data files *must* have the correct extension, however, the default extension can be changed in the so-called "expert edit mode" (see 2.3). Look for the field `fileExt=` and change the value to the desired extenstion.

Rather than entering the filenames manually, *PDFgetN* allows one to select the desired runs from a list. Simply use the *right* mouse button and click into the desired data file input field and a list will appear (see Figure 2.3). The dialog lists the file names as well as the run titles, making finding the correct file easy. Highlight the desired files and click *select*. Multiple files can be selected by holding down the *control* key when clicking on the files.

2.3 Action buttons

The bottom section of the main window contains all buttons needed to actually have *PDFgetN* do anything. This part is highlighted in Figure 2.4. On the left we find four buttons for process control. *Delete all* will delete all intermediate files. Unsually the processed characterization runs do not need to be reprocessed unless e.g. smoothing parameters have been changed. Using this button forces *PDFgetN* to reprocess all files. The three buttons next to it execute different external programs that actually process the data: *Create S(Q)* will combine and normalize the data and apply all corrections. The resulting $S(Q)$ is still separate for the individual banks. *Blend banks* will do exactly that and combine the values of the different detector banks to a single $S(Q)$. Finally *Create G(r)* will cut the blended data and calculate the resulting $G(r)$. As a step is completed, the corresponding button turns *green*.

The group on the right of botton section consists of a drop down menu and a *Do it* button. The philosophy is to select the desired action from the list and click on the button for execution. The choices can be seen in Figure 2.4 (right). The first two choices are to view or print the current plot (see section 2.7). Next the output

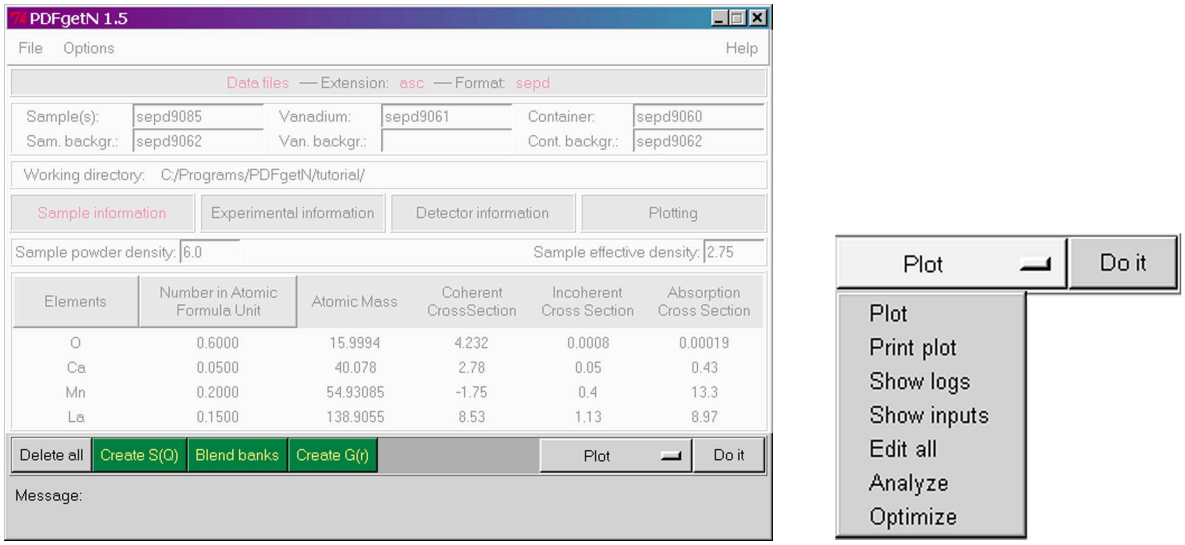

Figure 2.4: Processing control area (left) and drop down menu selections of *PDFgetN*.

and input files of every external processing program can be viewed. Seeing the input files is mainly useful for debugging, however, the output log files contain much interesting information about the processing run. The item *Edit all* starts the "expert edit" mode which allows one to directly modify the current history file since not all setting have a spot in the graphical users interface. A detailed discription of the history file is given in appendix A.3. The item *Analyze* lists various quality factors (Peterson et al., 2003b) for the current PDF. Finally *Optimize* allows one to automatically optimize the PDF as we will see in section 3.3.

2.4 Sample information

The sample information panel is shown in Figure 2.5. To activate it, simply click on the *Sample information* tab. On the top it contains the sample density and the effective sample density, ρ_{eff}. The sample density is not used in any calculation and can e.g. be used to record the bulk density of the material. The effective density, ρ_{eff}, however is a very important quantity controlling the normalization of $S(Q)$. *The simple rule is, if the effective density increases, the asymptote of $S(Q)$ decreases and vice versa*. The goal is to normalize $S(Q)$ at high values of Q to one. In many cases this is the only parameter that needs to be adjusted to get a good PDF. It is now also possible to let the program automatically determine the best effective density value. Simply click the *Automatic normalization* button next to the effective density field. The status line on the bottom of the main window will display information regarding the progress. Basically *PDFgetN* will process the data for the current density as well as for a slightly different density. Based on the average values of $S(Q)$ in both cases, the optimal density is calculated. Then the input field is updated and the final $S(Q)$ and $G(r)$ calculated. If the inital guess for the density is too far of, the process might fail or might need to be repeated.

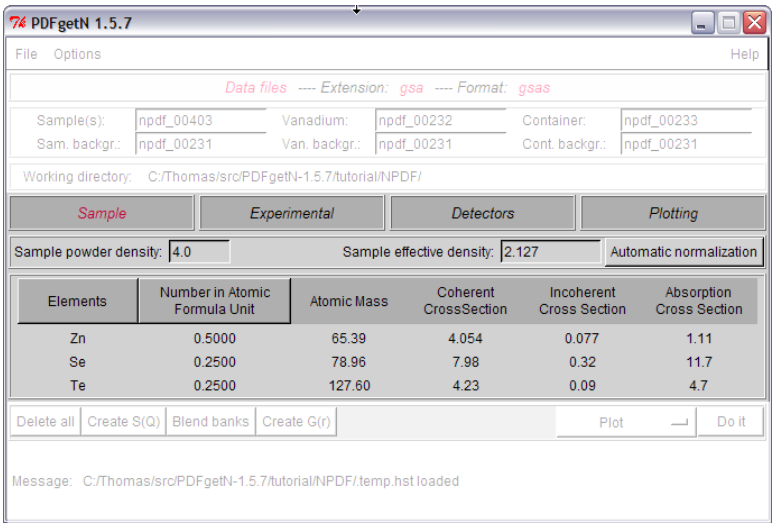

Figure 2.5: Sample information panel of *PDFgetN*.

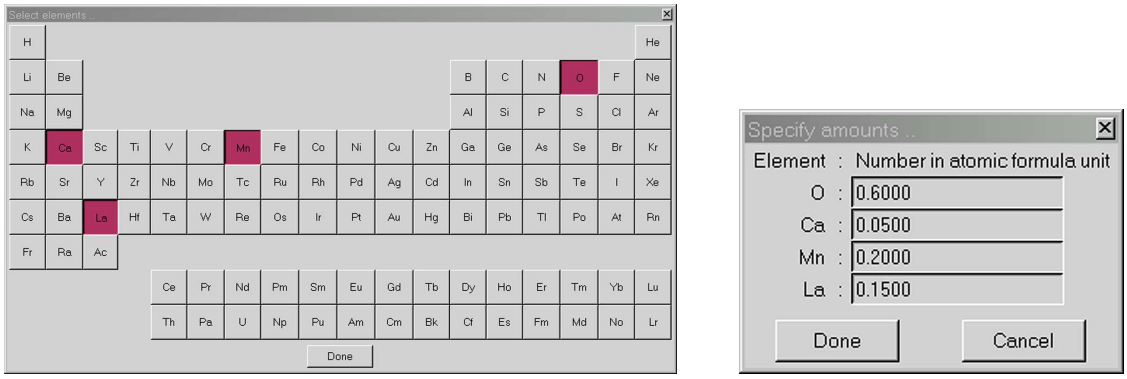

Figure 2.6: Element selection dialogs: Periodic table to select elements (left) and dialog to specify relative amounts (right).

Below the chemical composition of the sample is listed. To change or add elements, click on the button *Elements* and a periodic table will appear allowing one to select the desired elements (see Figure 2.6 left panel). Once the elements are selected, a second dialog appears (Figure 2.6 right panel) asking for the relative composition. Any notation works, however, internally the fractions are normalized to one and these normalized values appear in the table. All other values such as cross sections are taken from a lookup table. Note, that in order to use isotopes, one needs to enter the "expert edit mode" and provide all the information. The name of the element needs also to be changed, e.g. from H to D, otherwise the given values will be overwritten. Also if you want to update some of the cross sections e.g. with more recent values, you need to rename the elements in some way, since *PDFgetN* will overwrite all values of element names that match names in the lookup table.

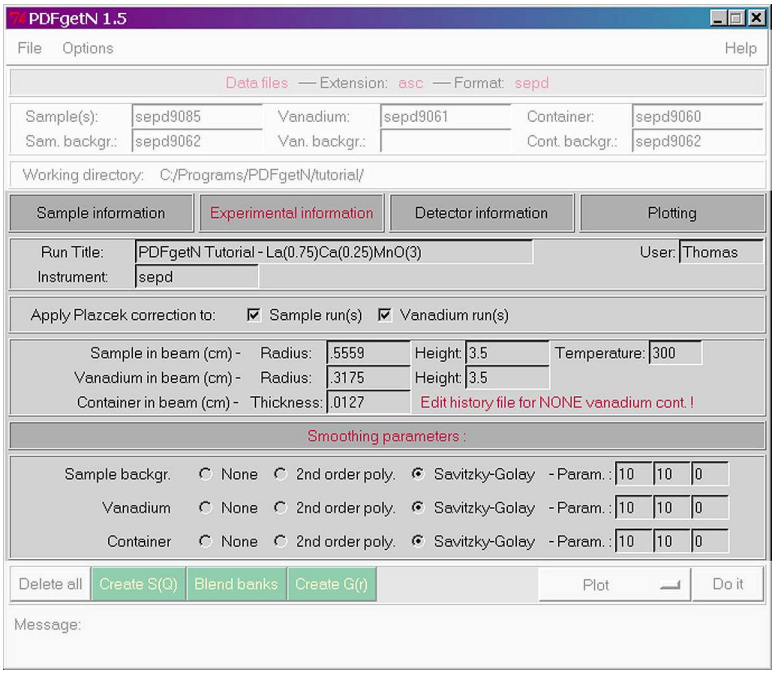

Figure 2.7: Experimental information panel of *PDFgetN*.

2.5 Experimental information

The experimental panel is shown in Figure 2.7. Again it is activated by clicking on the *Experimental information* button. Starting from the top we find the experiment title, user name and name of the instrument the data were collected on. Below one can select, if the Plazcek correction for the sample and vanadium can is applied or not. Further below are some geometric dimensions of the sample can, the vanadium rod and the beam height. All values are given in centimeters. Note that *PDFgetN* can only deal with upright cylindrical samples correctly. Also if a container material other than vanadium is used, some modifications in "expert edit mode" are required, since the cross section and density information for the container material need to be changed. The fields to be changed are in the section starting with ##### `Container` (see section A.3). The last section in the experimental panel deals with smoothing parameters for the characterization runs. Note that on case you are dealing with GSAS files, an additional field appears to specify the GSAS instrument parameter file corresponding to the measurement.

2.6 Detector bank information

Next we discuss the detector information panel and once again it is activated by clicking on the *Detector information* button. The panel consists of a table with detector bank information and below a section to specify information about the final $S(Q)$ and the PDF itself. The header of the detector table contains three

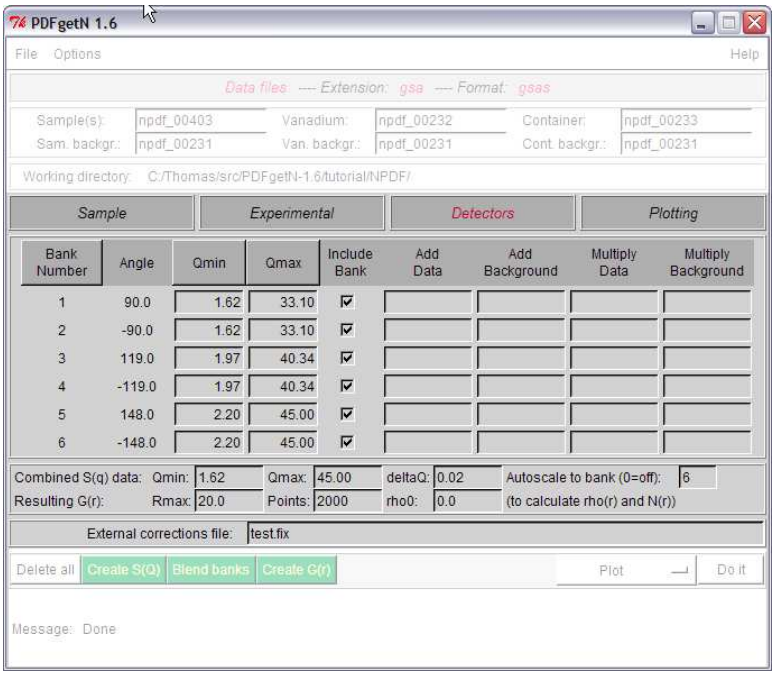

Figure 2.8: Detector bank information panel of *PDFgetN*.

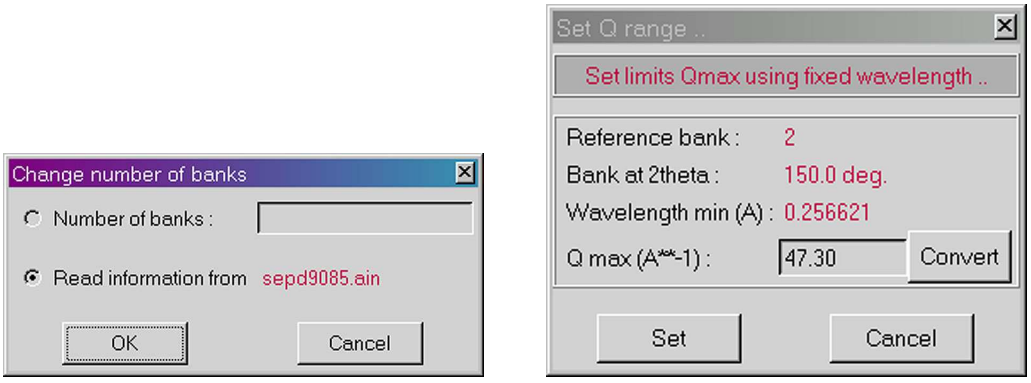

Figure 2.9: Detector bank dialogs: Selection of number of banks (left) and constant wavelength limits dialog (right).

buttons: Clicking on *Bank Number* invokes the dialog box shown in Figure 2.9 (left). One can either specify the number of banks corresponding to the current data set and enter all other information manually or read the information from a preprocessed intermediate file. In this case bank angles and Q range are read from the file. In case no banks are defined, the processing will top after that intermediate file is produced and the bank information can be imported. However, the values in columns labelled Q_{max} and Q_{min} can be changed to the desired limits for the blending step. In many cases data contain spurious features at the boundaries that need to be excluded. As an alternative one can use a Q range for each bank that corresponds

to a constant range in neutron wavelength, λ. To specify the lower or upper Q limits corresponding to a given wavelength, click on the *Qmin* or *Qmax* buttons. The dialog box shown in Figure 2.9 (right) will appear. Simply enter the highest/lowest Q value and click the *convert* button. You see the wavelength value update and by then clicking the *set* button, the Q values corresponding to the displayed wavelength will be set. Column *Include Bank* allows one to include or exclude a complete bank from the blending process. The last four columns contain possible correction factors or offsets for the sample or background data. Note that these values are applied to the *raw* data before normalization.

After the data from the different banks are blended into a single $S(Q)$, the final data need to be truncated and rebinned on a grid with constant ΔQ. The top row below the detector bank information contains the relevant information. One can specify the minimum and maximum Q value to be used for the final $S(Q)$. The program extrapolates $S(Q)$ from the given lower limit to $Q = 0$. As default, the data are simply cut at the given upper limit, *PDFgetN* however, allows to apply a damping function as well. This again needs to ge specified in "expert edit" mode and for details about the available functions, refer to appendix A.3. Note, that the modelling programs *PDFFIT* and *DISCUS* can only model the effect of termination for a simple cut. The last field in that row contains a reference bank number. If a valid bank number is give, the $S(Q)$ values of the different banks will be modified to give the best overlap possible by automatically applying a scale and offset. If the reference bank is set to zero, no such modification is carried out. The input fields in the bottom row of the detector information panel specify the maximum value r_{max} and the number of points, N, for $G(r)$. Note that currently *PDFgetN* will always calculate the PDF from Δr and the step width is given by $\Delta r = r_{\mathrm{max}}/N$. The last value given in that row is the number density ρ_0. If a nonzero value is given, *PDFgetN* will calculate the functions $\rho(r)$ (Eq. 2.1) and the coordination number $N(r)$ (Eq. 2.2) as well as $G(r)$.

$$\rho(r) = G(r)/(4\pi r) + \rho_0 \tag{2.1}$$

$$N(r) = \int_0^r 4\pi r^2 \rho(r) dr \tag{2.2}$$

The bottom filed on this panel allows one to specify a file containing scale factors and offsets for each point Q. These corrections are applied to the individual $S(Q)$ data sets for each bank just before merging. Details are discussed in section 3.4. As for the other fields, right clicking on the entry field will bring up a file selector box.

2.7 Plotting

PDFgetN has a build-in plotting capability. To switch to the plotting controls panel, simply click on the *Plotting* button. Note that *PDFgetN* requires the plotting program *KUPLOT* for the actual plotting. If *KUPLOT* is not found, the panel will contain just the message *KUPLOT not found*. Under Windows, *KUPLOT* is included in the binary distribution and this problem should not occur. If you are using Linux/Unix however, *KUPLOT* needs to be installed separately and the binary must be in the path, in other words, if you type `kuplot` the program should start. Details can be found in appendix B.2.

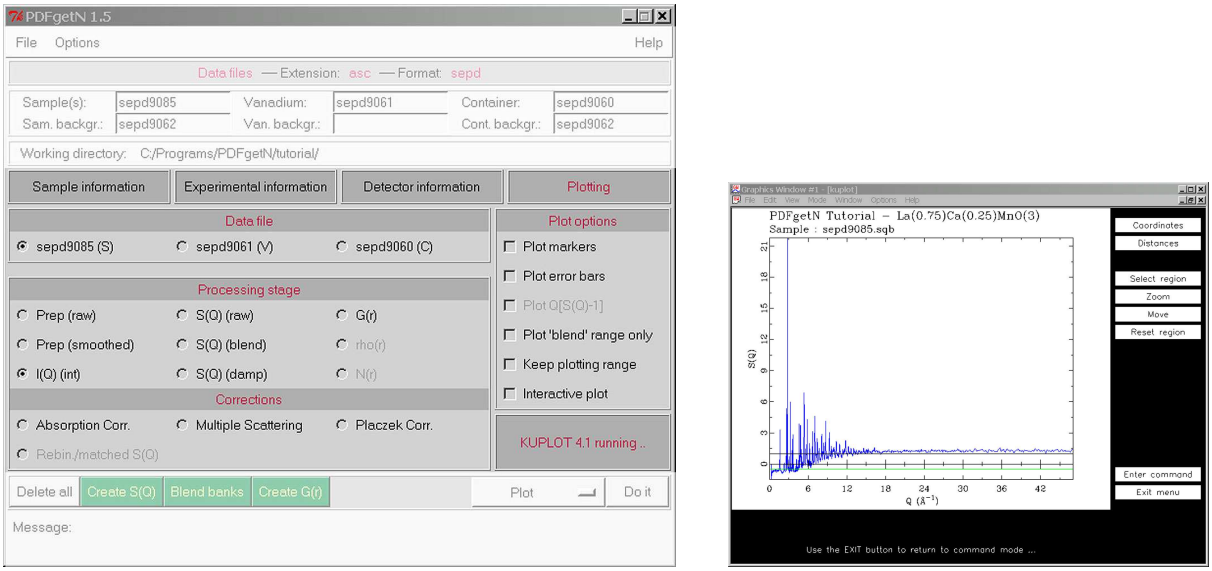

Figure 2.10: Plotting control panel (left) and plotting window (right) of *PDFgetN*.

Processing stage	**What is it ?**	**Data**
Prep (raw)	Raw data; if multiple files are given, they are merged at this stage.	B,V,C
Prep (smoothed)	Same as above, just after smoothing is applied.	B,V,C
I(Q) (int)	Intensities after normalization and background subtraction.	S,V,C
S(Q) (raw)	$S(Q)$ after corrections are applied, but still separate for individual banks.	S
S(Q) (blend)	$S(Q)$ after merging of all banks and rebinning on constant Q grid.	S
S(Q) (damp)	$S(Q)$ after final termination and extrapolation to $Q = 0$.	S
G(r)	PDF $G(r)$	S
ρ(r)	$\rho(r)$ (if ρ_0 was specified)	S
N(r)	$N(r)$ (if ρ_0 was specified)	S

Table 2.1: List of processing stages for plotting.

If all works as it should, the plotting controls panel shown in Figure 2.10 on the left will appear. The buttons on the left allow one to select what data will be plotted. All the intermediate files are ASCII and can be plotted using other programs as well. Details about file names and formats are given in appendix A.2. In section labelled *Data file* one specifies if sample, vanadium or container data should be plotted. Of course after processing step $S(Q)$ and beyond, only sample data are available and the other two buttons become disabled. In section *Processing stage* one specifies which intermediate file is to be plotted. A listing and explanation of the different choices is given in Table 2.1. The bottom left area *Corrections* allows one to select the absorption, multiple scattering and Placzek correction to be plotted. If automatic bank matching was selected (see section 2.6), one can also choose to plot the rebinned and rescaled $S(Q)$ data.

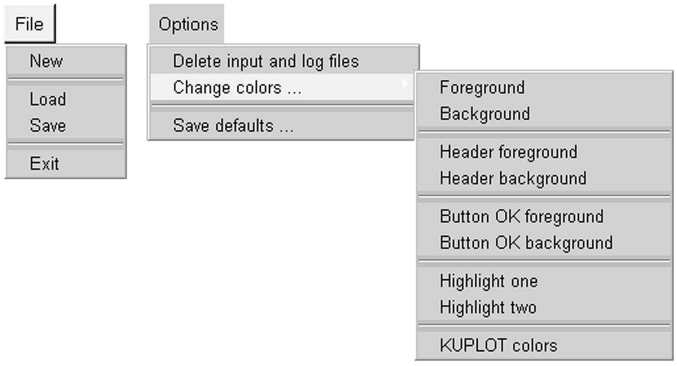

Figure 2.11: Drop down menu of *PDFgetN*.

The checkboxes on the right of the plotting controls panel are used to specify certain plotting options, such as plotting markers or error bars. If *Plot 'blend' range only*, the range in Q that is displayed is truncated according to the values given in the detector information panel. Otherwise, the full range present in the data is shown. If *interactive plot* is checked, the plot window will contain several buttons allowing one to specify a region of interest and so one. By clicking in the button *Do it* with *Plot* selected in the action buttons area of *PDFgetN*, the selected data will be plotted (see Figure 2.10 right). Note, that in case of an "interactive" plot, one needs to exit the interactive mode by clicking the *Exit menu* button to regain control of the main *PDFgetN* window. In order to print (or save) the current plot, simple select *Print Plot* from the menu next to the *Do it* button and hit the *Do it* button. If you are familiar with *KUPLOT* commands, you can use the *Enter Command* button on the plotting screen to enter any valid *KUPLOT* command. However, due to the current implementation one needs to type *really slow* to avoid characters being lost. Finally, the color scheme of *KUPLOT* can be customized using the *Options* menu as discussed in section 2.8. If you have followed along using the tutorial data and plot the final $S(Q)$, you will notice, that the data are not quite normalized at high Q. The horizontal line in the plot marks $S(Q) = 1$. As a practice, you should try adjusting the effective density and reprocess the data until you have a good normalization.

2.8 PDFgetN menu

The only part of the *PDFgetN* graphical users interface, we have not discussed so far are the menus which are shown in Figure 2.11. First let us look at the *File* menu: The choice *New* will bring up the startup dialog (see 2.1) and one can start a new data processing session. The entries *Save* and *Load* will save or load a history file. The three files that are saved are the final $S(Q)$ (file.sq)m the PDF $G(r)$ (file.gr) and the processing information (file.hst). Note that the $S(Q)$ as well as $G(r)$ data files have the same processing information pre-pended to the data, so any of the three files can be loaded into *PDFgetN*. The entry *Save GSAS* allows one to save the processed $S(Q)$ values in GSAS file format. This allows one to use the corrected $S(Q)$ data as input for Rietveld refinements. The entry *Exit* will (no surprise here) terminate the program. The menu

Parameter	What is it ?	Default
file	*PDFgetN* history input file.	
-gui	Start *PDFgetN* with graphical interface.	D
-nogui	Start *PDFgetN* without graphical users interface. If no filename is given the program will prompt for one.	
-keep	Keep intermediate files.	
-nokeep	Delete intermediate files.	D
-opt	Run automatic optimization on final $G(r)$ (see 3.3).	
-noopt	No automatic optimization.	D
-norm	Run automatic normalization on $S(Q)$.	
-nonorm	No automatic normalization.	D
-Ppar=val,..	Overwrite values for processing parameters (see text)	
-help	Print help message.	
-debug	Run the interface without actually processing data.	

Table 2.2: *PDFgetN* command line parameters.

entry *Options* has three entries. If the first choice *Delete input and log files* is checked, *PDFgetN* will delete all intermediate files upon exiting the program, otherwise all intermediate files will remain after the program is terminated. The next entry, *Change colors* .., allows one to change the colors of the *PDFgetN* interface as well as the *KUPLOT* plotting colors. The new colors selected to the *PDFgetN* interface will only be shown after the defaults have been saved and the program has been restarted. The last entry, *Save defaults* .., will save the current *PDFgetN* settings as new default values.

2.9 Batch processing

So far we have used *PDFgetN* through its graphical users interface. This is the default behavior if no command line parameters are given. However, *PDFgetN* can also be used for batch processing of data. In this case the program needs to be started from the command line with one or more command line parameters. A list of valid parameters is given in Table 2.2. The command line parameter `-P` allows one to overwrite certain values of processing parameters. The valid parameters are `run` for the data file name, `dens` for the effective density and `temp` for the temperature. This way one can create scripts e.g. in *PERL* that modify a certain value in the history file and run a series of *PDFgetN* processing runs. The following example will use the history file `rt.hst` containing all processing parameters obtained e.g. from the room temperature data set and process $G(r)$ and $S(Q)$ for two subsequent runs measured at different temperatures by changing the run number and temperature as well as performing an automatic normalization:

```
PDFgetN -nogui -norm -Prun=data_100K,temp=100 rt.hst
PDFgetN -nogui -norm -Prun=data_50K,temp=50 rt.hst
```

Chapter 3

Optimizing the PDF

In this chapter we will give a *very* short overview about experimental considerations for a neutron PDF measurement as well as some hints how to optimize the resulting $G(r)$ using *PDFgetN*. A more detailed summary about the theoretical PDF data processing details can be found in Toby and Egami (1992) and in the new book by Egami and Billinge (2003). For an even more theoretical approach, one might refer to the book by Warren (1990).

3.1 Experimental considerations

In this section we will discuss a few basic considerations regarding the neutron scattering experiment itself. First of all, because the PDF is obtained via Fourier transform from the scattering data, the extend of the data in Q-space determines our real space resolution of the PDF. This is a result of the *termination effect* which manifests itself in a convolution of the data with the Fourier transform of the cut-off function of $S(Q)$. In case of a simple cut, this function is simply $\sin(Q_{\text{max}}r)/r$. The effect of Q_{max} is illustrated in Figure 3.1. Here we see the nearest neighbor (NN) peak of a semiconductor alloy (Peterson et al., 2001). The solid line corresponds to a termination at $Q_{\text{max}} = 40\text{Å}^{-1}$, the highest value achievable for these data. One can clearly see that the NN peak is split, in fact we observe two distinct bond length showing that *locally* the structure retains the end member NN bond length. The dashed line in Figure 3.1 shows the same data, but this time terminated at$Q_{\text{max}} = 17\text{Å}^{-1}$, a value corresponding to MoKα radiation or a spectrum collected at a reactor neutron source. In this case we do *not* observe the peak split. From this it is apparent that the required real space resolution, Δr will determine the value of Q_{max} required. However, using pulsed neutrons or high energy X-rays, values of $Q_{\text{max}} > 50\text{Å}^{-1}$ can easily be achieved.

The influence of the Q resolution, ΔQ, on the PDF can easily be seen by inspecting Fig. 3.2 showing $G(r)$ for Ni obtained from the new NPDF diffractometer at the Lujan Center and the GLAD instrument at the Intense Pulsed Neutron Source. In all cases the data were processed using *PDFgetN* and the scattering data for both instruments were terminated at $Q_{\text{max}} = 35\text{Å}^{-1}$. Note that the PDF in Fig. 3.2 is shown up to distances $r = 100\text{Å}$. The Q resolution results in an exponential dampening of the PDF peaks as function

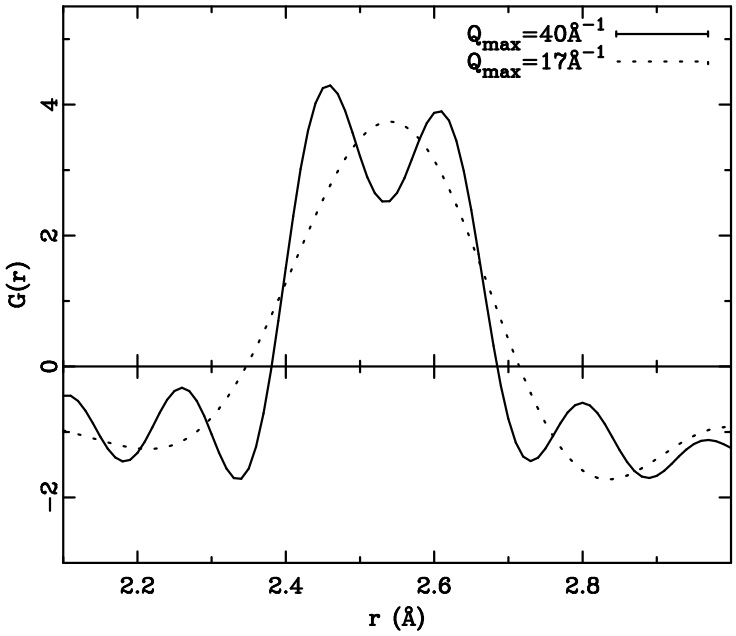

Figure 3.1: Nearest neighbor peak of PDF of semiconductor alloy with $S(Q)$ terminated at $Q_{\max} = 40\text{Å}^{-1}$ (solid line) and $Q_{\max} = 17\text{Å}^{-1}$ (dashed line). Data from Peterson et al. (2001).

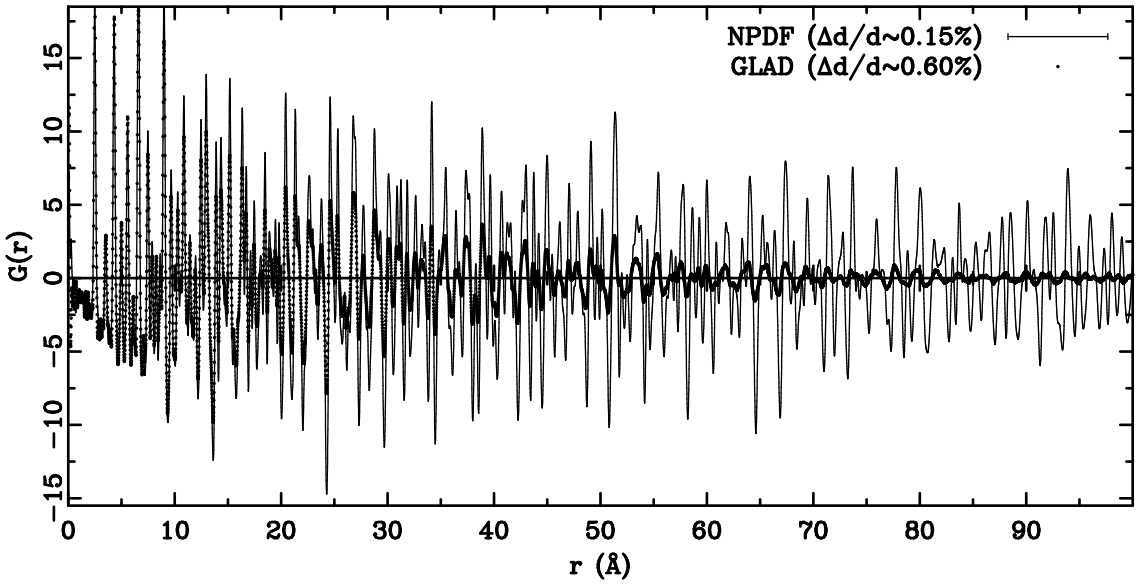

Figure 3.2: Influence of the reciprocal space resolution ΔQ on the PDF. The solid line are data from NPDF at Lujan Center at Los Alamos National Laboratory and the dots are data from GLAD at Intense Pulsed Neutron Source at Argonne National Laboratory (from Proffen et al. (2002)).

Facility	Instrument	Specifications
Lujan Center	NPDF	Resolution: $\Delta d/d \approx 0.15 - 0.31\%$, Q range: $\approx 1.5 - 51.1\text{Å}^{-1}$ `http://lansce.lanl.gov/lujan/instruments/npd/npd.htm`
	HIPPO	Resolution: $\Delta d/d \approx 0.40 - 5.0\%$, Q range: $\approx 0.13 - 52.4\text{Å}^{-1}$ `http://lansce.lanl.gov/lujan/instruments/hippo/hippo.htm`
	HIPD	Resolution: $\Delta d/d \approx 0.30 - 3.0\%$, Q range: $\approx 0.2 - 60.0\text{Å}^{-1}$ **(not in users program)** `http://lansce.lanl.gov/lujan/instruments/hipd/hipd.htm`
IPNS	GPPD	Resolution: $\Delta d/d \approx 0.26 - 0.76\%$, Q range: $\approx 1.1 - 60.4\text{Å}^{-1}$ `http://www.pns.anl.gov/instruments/gppd/`
	SEPD	Resolution: $\Delta d/d \approx 0.34 - 4.7\%$, Q range: $\approx 0.6 - 47.3\text{Å}^{-1}$ `http://www.pns.anl.gov/instruments/sepd/`
	GLAD	Resolution: $\Delta d/d \approx 0.6 - 5.0\%$, Q range: $\approx 0.2 - 45.0\text{Å}^{-1}$ `http://www.pns.anl.gov/instruments/glad/`
ISIS	GEM	Resolution: $\Delta d/d \approx 0.34 - 4.7\%$, Q range: $\approx 0.6 - 100.0\text{Å}^{-1}$ `http://www.isis.rl.ac.uk/disordered/gem/gem_home.htm`
	POLARIS	Resolution: $\Delta d/d \approx 0.5 - 1.0\%$, Q range: $\approx 0.3 - 31.5\text{Å}^{-1}$ `http://www.isis.rl.ac.uk/crystallography/polaris/`

Table 3.1: List of neutron powder diffractometers for PDF studies at Lujan Center at Los Alamos National Laboratory, USA, Intense Pulsed Neutron Source (IPNS) at Argonne National Laboratory, USA and at ISIS at Rutherford Appleton Laboratory, England. The information is taken from the respective instruments WWW pages.

of r (Toby and Egami, 1992). Using high Q-resolution data for PDF analysis from an instrument such as NPDF, allows one to access medium range distances, $r \approx 10 - 100$Å, opening up a new territory of research. In summary, the required real space resolution of the PDF will determine the necessary Q_{max} and the desired maximum range in r will determine the Q space resolution ΔQ one needs. One other question is what would be considered sufficient statistics for a PDF measurement. Note that the overall counting time will be dominated by the count rate at high Q values. One rule of thumb is the so-called *Egami rule* which states that one wants to accumulate 10^6 *neutrons / Å*$^{-1}$ in the high Q region of the measurement (Toby and Egami, 1992). Another very important consideration is to choose an instrument and ancillary equipment with a very low background scattering, since we are not only interested in Bragg scattering but also in the weak diffuse scattering. A list of current neutron powder diffractometers, the authors have successfully used to measure PDF data are listed in Table 3.1. Finally a PDF measurement usually requires several characterization runs in addition to the sample measurement. On some instruments standard characterization runs are measured before the run cycle starts and are already available, on other instruments or when using special ancillary equipment, those measurements become part of the experiment. Usually three measurements besides the sample run are required: *vanadium*, *empty container* and *instrument*

background. In some cases the background might be so low, that a background measurement is not really necessary. In some cases the vanadium runs was performed using a different setup, so it needs its own background run. However, it is recommended, to carry out all runs in the same instrument configuration.

3.2 Normalization

One of the most important tasks when processing data to $S(Q)$ is to obtain a correct normalization, i.e. $S(Q) \rightarrow 1$ for $Q \rightarrow Q_{\text{max}}$. Finding the correct *single scattering* $S(Q)$ means having the correct sample container dimensions, beam dimensions and the effective sample density ρ_{eff}. Unfortunately the $S(Q)$ obtained can be distorted by a scale or offset or both and all the required information listed above has its own error bars. Usually one uses the effective sample density ρ_{eff} as the "knob" to obtain the proper normalization. If ρ_{eff} is increase, the resulting $S(Q)$ will decrease and vice versa. However, the density can compensate for errors e.g. in the beam dimensions resulting in a scaled PDF. For structural refinements using e.g. *PDFFIT* a build in scale factor will take care of the problem and its deviation from unity will be the indication how good the normalization has been. If one, on the other hand, wants to obtain coordination numbers by integrating PDF peaks, a scale factor is unacceptable. In those cases, a reference sample with known coordination numbers can be used, to obtain the correct sample and beam dimensions and those are then used for subsequent data processing. In any case a properly normalized $S(Q)$ is a requirement to obtain a meaningful PDF $G(r)$. Most of the time visual inspection of $S(Q)$ using the build in plotting function of *PDFgetN* will be sufficient to judge the $S(Q) \rightarrow 1$ criterion. However, the program also calculates the average $< S(Q) >$ over a given Q range. The value is listed in the analyze dialog (see section 2.3).

As a new feature in version 1.6, the program can determine the effective sample density needed to fullfill $S(Q) \rightarrow 1$. This is simply done by exploiting the relation ship between the density and $S(Q)$ (see equation (5.27) in Egami and Billinge (2003)). The program obtains the average value $\langle S(Q) \rangle$ for two different densities and extrapolates the required density for $\langle S(Q) \rangle = 1$. Note that automatic normalization can be achieved via this option adjusting the density or via the automatic optimization tool discussed in section 3.3. The main difference is that changing the effective density not only scales $S(Q)$, but also changes certain corrections such as absorption or multiple scattering. One should take extra care to ensure that the density value used makes physical sense. The automatic optimization on the other hand only applies a scale and offset to the final values of $S(Q)$. One can also use both options to obtain a sensible density and further optimize the PDF to minimize ripples at low values of r (see 3.3).

3.3 Automatic optimization

The ultimate feature of any processing program is the automatic, reliable optimization of the processing process. *PDFgetN* has the capability of optimizing the PDF by changing the scale and offset of the final

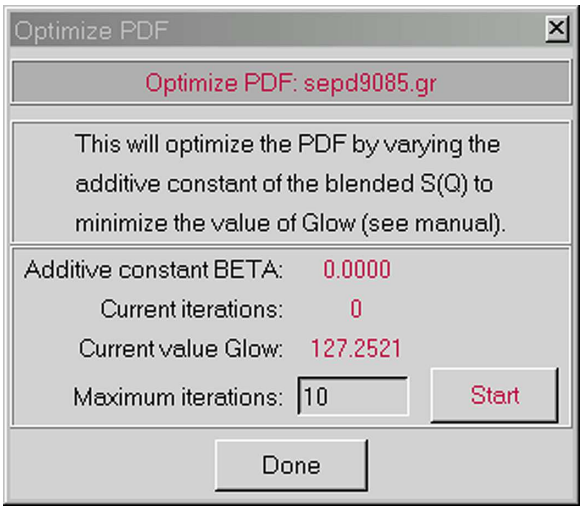

Figure 3.3: Auto-optimization dialog.

$S(Q)$ in order to minimize G_{low} defined as

$$G_{\mathrm{low}} = \frac{\int_0^{r_{\mathrm{low}}} [rG(r) + 4\pi r^2 \rho_{\mathrm{fit}}]^2 dr}{\int_0^{r_{\mathrm{low}}} [4\pi r^2 \rho_{\mathrm{fit}}]^2 dr}. \tag{3.1}$$

Here the integral runs up to r_{low} which is a value *below* the first PDF peak and ρ_{fit} is a fitted number density. In general G_{low} is a measure for the noise ripples below the first PDF peak. Details about these equations as well as the optimization procedure are given in Peterson et al. (2003b). In this section we just focus on the implementation in *PDFgetN*.

The automatic optimization is started by selecting *optimize* from the action menu and clicking the *Do it* button (see 2.3). This will bring up the dialog box shown in Figure 3.3. The dialog lists the current additive constant β, the number of iterations so far as well as the current value of G_{low} discussed above. The last line allows to specify the maximum number of iterations *PDFgetN* is allowed to attempt before terminating. The optimization process is started by clicking the button *Start*. The program will modify $S(Q)$ according to

$$S'(Q) = \alpha S(Q) + \beta. \tag{3.2}$$

The parameter β is modified and α is recalculated obtain the best normalization $S(Q) \rightarrow 1$ for $Q \rightarrow Q_{\mathrm{max}}$. In every iteration the new value of β and G_{low} are displayed. Once the shift in β becomes very small or the maximum number of iterations is reached, the process will stop. TO exit the dialog, simply click the *Done* button. The PDF will then be the optimized one and can be saved the usual way. Please note, that this feature is in its early stages and might not work as expected on certain data. The value r_{low} is specified in the history file in the entry `intMaxR=`. It can be changed using the "expert edit" mode (see 2.3). Again, much more detailed information about the optimization strategy can be found in Peterson et al. (2003b).

3.4 Generic corrections

In some cases one might be faced with certain undesired features in $S(Q)$ that can not be corrected using the standard methods available in *PDFgetN*. The most common case might be dealing with the incoherent scattering contribution of e.g. hydrogen. In these cases an external program might be used to obtain a correction function e.g. by fitting a polynomial to the data. These corrections can be put back into *PDFgetN* using a correction file which can be specified in the detector information panel of the program. Usually one would take the rebinned $S(Q)$ values stored in the intermediate file `blen_bin.dat` as a basis and create a correction file on a similar grid in Q. It is not required to match the Q grid exactly, since the correction data are rebinned by the program *PDFgetN* before they are applied. The required file format for the correction file is as follows:

```
#S 1   Correction for bank 1
#L Q   MUL ADD
 1.00   1.00 0.7654
 1.02   1.00 0.7865
  :      :    :
45.00   1.00 0.5433
#S 2  Correction for bank 2
#L Q   MUL ADD
 1.00   1.00 1.7654
 1.02   1.00 1.7865
  :      :    :
```

A few notes about correction files: The line starting with `#S n` specifies the start of the data section for bank `n`. *Each bank must have an entry in the file.* The line starting with `#L` is optional, but including it makes this a valid *SPEC* file that can be plotted using e.g. *KUPLOT*. After the header lines, for each point in Q, the following three numbers a specified in a single line: value of Q_i, factor α and offset β to be allied to the value of $S(Q_i)$.

Appendix A

File formats

A.1 Supported data file formats

In this section, all data file formats *PDFgetN* can read are summarized. However, it is likely that future versions of *PDFgetN* will *only* support the use of GSAS files which can be obtained at most neutron powder diffractometer. For this reason, the use of the other file formats for new measurements is not recommended. Also, some of the other formats are binary in nature and need to be converted to an ASCII format using utilities distributed with the *PDFgetN* package.

GSAS files

GSAS refers to the General Structure Analysis System (Larson and Von Dreele, 2000). The GSAS file format is keyword driven and can be used to describe neutron as well as X-ray data. Obviously *PDFgetN* can only read a subset of valid GSAS files. *The data must be time-of-flight (TOF) neutron scattering data using either TIME_MAP, CONS or FXYE as binning scheme.* In addition an GSAS instrument parameter file is required, containing calibration information in particular diffractometer constants to convert from TOF to d-spacing. However, most powder diffractometers located at a spallation neutron source will produce the correct file and the instrument scientist can usually provide the correct instrument parameter file. In short if you can use GSAS for a Rietveld refinement of you data, you can generally use *PDFgetN* to read them as well. There is one problem that might occur, not all GSAS files have a `Monitor` record which is required by *PDFgetN* to properly normalize the data. If this is the case, a simple modification of the ASCII GSAS file is required. The monitor count needs to be added as line 3 (see below).

```
DISPLEX V-NB (4583)
Instrument parameterfile:[ANALYSIS.CALIBRATION]IPARM_11JUN01.NPD
Monitor: 1597078
   :
```

Note that the monitor count can be anything measuring the amount of neutrons received during the run, e.g. proton current, monitor count, pulses received, as long as the same measure is used in all files processed together.

Some notes on GSAS file produced on different instruments: The program `hdfbin` used on HIPPO at the Lujan Center, the question "*Normalize by monitor ?*" needs to be answered with *Y*. This will produce a GSAS file compatible with *PDFgetN* even though it contains no line with `Monitor`. On NPDF at the Lujan Center, the resulting GSAS files are already normalized by the beam monitor counts and monitor line reads: `Monitor: 1`. Finally all GSAS files produced by the Integrated Spectral Analysis Workbench software (Peterson et al., 2003a) developed at IPNS has the option to include the integrated monitor in the resulting GSAS file.

IPNS run files

For IPNS run files collected on SEPD and GPPD there exists a direct binary-to-ASCII conversion program called sepdbtoa. The raw data files as they appear from the data collection software have to be copied to your computer via binary ftp. The usage of sepdbtoa is as follows:

```
sepdbtoa -h 1 file1.run file2.run file3.run...
```

where file1, file2 etc. are the filenames of the corresponding run files. They consist typically of the instrument name and run number, e.g. SEPD02258 or GPPD01157. The parameter `-h 1` specifies the histogram to be extracted, usually histogram 1, which is the default. Typing `sepdbtoa` without specifying any filename will also remind you of its usage. An utility program called `asc2plot` can be used to convert raw ASCII files to the SPEC file format (SPEC, 2003) that can be read by *KUPLOT*. The usage of `asc2plot` is as follows:

```
asc2plot sepd12345.asc sepd12345.plot
```

where the first file is the ASCII file (.asc) which is then converted to the output file given as the second argument. Any suffix or file name can be used for the output file. For information how to use *KUPLOT* to plot these files, refer to the *KUPLOT Users Guide*.

ISIS norm files

At some facilities the data can be preprocessed locally to obtain a normalized intensity file versus Q. This is done on GLAD (IPNS) as well as LAD and POLARIS (ISIS) using NORM. The files are binary GENIE intermediate files usually with extension .nrm. These files need to be copied to your computer via binary ftp and then need to be converted to ASCII. At the time of writing of this manual no direct conversion program exists. However, the program OPEN GENIE is capable of performing this step. This program is freely downloadable at `http://www.isis.rl.ac.uk/OpenGENIE/` and is available for a variety of platforms. The *PDFgetN* distribution contains the script 'nrmtoasc.gcl' in the directory 'raw2asc/isis'. To use the macro one can load it within OPEN GENIE using the command:

```
>> load "nrmtoasc.gcl"
```

where the >> is the OPEN GENIE prompt and need not to be typed. The conversion is then performed using the command:

```
>> nrmtoasc "inst_name" first_run last_run number_of_banks
```

where `inst_name` is a four-letter shortcut for the name of the diffractometer. Note that the " have to be written explicitly. `First_run` and `last_run` are the run numbers of the first and last run to be converted, respectively, without suffix. This allows the user to convert a range of files. `Number_of_banks` is the number of banks of the instrument. Note that there is a bank 0 which is identical to bank 1, i.e. the data of bank 1 appears two times, namely at bank 0 and bank 1. Thus, supply for `number_of_banks` the actual number of banks for a given instrument adding one otherwise the last bank will be missing. Example:

```
>> nrmtoasc "lad" 14933 14936 8
```

In this case, NORM-files from LAD are used. The run numbers from 14933 to 14936 will be converted to ASCII. LAD has seven banks and thus `number_of_banks` is eight in this example. The resulting files will be called lad14933.asc etc.

ARIEL files

The diffractometer GEM at ISIS uses the program *ARIEL* for data reduction. This program produces an ASCII SPEC (SPEC, 2003) file with wavelength versus counts. These files can not only be plotted with *KUPLOT* directly but also be read by *PDFgetN*. Apart from processing GEM data, this file format might be the most straight forward and be used to convert data from a unsupported format to be used with *PDFgetN*. The header and first few data lines are shown below.

```
#F  H:\users\001\billinge_11598\gem01487.asc
#D  created Thu Jun 29 03:52:37 2000
#C  Instrument: gem
#C  Run GEM01487
#C  Title Vanadium 8mm - candlesti27-JUN-2000 19:52:27
#C  User SJB/TP/PGR
#C  Primary flight path 17.0000
#O0 lt0  lstep  angle  fpath
#S  1 - Group 3
#P0 6.05640 0.000799686 17.9541 18.7750
#L  W Data Error
    0.0899953  9.3768566e-006  2.0967288e-006
    0.0900673  1.1711701e-005  2.3423402e-006
    0.0901393  8.8937786e-006  2.0403727e-006
     :          :                :
```

The only lines of the header that are relevant to *PDFgetN* are explained below. First the line starting with `#O0` specifies labels for the parameter section and can just be copied. The corresponding line starting with `#P0` contains the corresponding values: logarithmic time binning constants T_0 (`lt0`) and ΔT (`lstep`). The time-of-flight for a point i is given by

$$TOF(N) = \exp(T_0 + N\Delta T). \tag{A.1}$$

The last two values in the header row starting with `#P0` are the bank angle 2Θ in degrees (`angle`) and the total flight path in meters (`fpath`). Note that only these last two values are used by *PDFgetN* since the data are already converted to wavelength. Next, the line starting with `#S` marks the start of a data block for a given bank, here group three. The line `#L` specifies the labels for the data columns: wavelength, intensity and standard deviation of the intensity. After this line, the data start until the next banks (`#S`) is encountered.

A.2 Intermediate files

As we mentioned before, *PDFgetN* is fact a front-end to the GLASS package and the different parts of that package interact via *binary* files. In addition *PDFgetN* also produces an ASCII file at every intermediate step, suitable for plotting of further analysis outside of *PDFgetN*. A list of all the intermediate files is given in Table A.2. If only a suffix is given, the filename will be the name of the sample data file with the given extension. Some files containing correction factors have a fixed name. Note that not for all ASCII files there is a binary analog. *All* ASCII intermediate files are written in SPEC format (SPEC, 2003) and can be read and plotted by e.g. *KUPLOT*. The files that contain the final $S(Q)$ (.sq) and final $G(r)$ (.gr) have the processing parameters pre-pended to the data. The start of the data themselves is marked by a line `##### start data`. The data section of a .sq file typically looks as follows:

```
##### start data
#S   1 Bank at    0.00 degrees
#L Q S sigmaS ReducedS sigmaReducedS
  47.2899971  2.85515881  0.0710189641  87.7304535  3.35848665
  47.2700081  2.85515881  0.0710189641  87.6933746  3.35706711
  47.2500038  2.85515881  0.0710189641  87.6562576  3.35564637
  47.2300034  2.85515881  0.0710189641  87.6191559  3.35422587
```

The columns are Q, $S(Q)$, $\Delta S(Q)$, $Q(S(Q)-1)$ and standard deviation of the latter. As mentioned, the PDF also contains the processing parameters pre-pended to it. The data section typically looks as follows:

```
##### start data
#O0 rg_int sig_rg_int low_int sig_low_int rmax rhofit
#S 1 - PDF from PDFgetN
#P0  -79.99439    0.50518    0.13495    0.00140    1.50  0.1054
#L r G(r) dr dG(r)
```

Suffix ASCII	**Suffix binary**	**Created by clicking on**	**Description**
.ain	.int	Create $S(Q)$	Normalized and background subtracted data.
.braw		Create $S(Q)$	Raw background data.
.bsmo		Create $S(Q)$	Smoothed background data.
.craw		Create $S(Q)$	Raw sample container data.
.csmo		Create $S(Q)$	Smoothed sample container data.
.vraw		Create $S(Q)$	Raw vanadium data.
.vsmo		Create $S(Q)$	Smoothed vanadium data.
soqd_cabs.dat		Create $S(Q)$	Absorption correction factors for sample container.
soqd_sabs.dat		Create $S(Q)$	Absorption correction factors for sample.
soqd_smsc.dat		Create $S(Q)$	Multiple scattering correction factors for sample.
soqd_splc.dat		Create $S(Q)$	Placzek correction for sample.
soqd_vabs.dat		Create $S(Q)$	Absorption correction factors for vanadium.
soqd_vmsc.dat		Create $S(Q)$	Multiple scattering correction factors for vanadium.
soqd_vplc.dat		Create $S(Q)$	Placzek correction for vanadium.
.sqa	.soq	Create $S(Q)$	Corrected final $S(Q)$, still bank by bank.
blen_bin.dat		Blend banks	Rebinned and rescaled (if selected) $S(Q)$, still bank by bank.
.sqb	.bld	Blend banks	Merged $S(Q)$, full range.
.sq	.asq	Create $G(r)$	Final, truncated $S(Q)$.
.gr	.pdf	Create $G(r)$	Pair distribution function $G(r)$.

Table A.1: List of intermediate files produced by *PDFgetN*.

```
0.010      1.643   0.0     0.0049
0.020      3.199   0.0     0.0093
0.030      4.585   0.0     0.0130
```

The columns are distance r in Å, $G(r)$ in Å^{-2}, Δr in Å, and $\Delta G(r)$ in Å^{-2}, where $G(r)$ is the pair-distribution function. If the number density of the sample is given in the history file then the file will also contain the pair-density function $\rho(r)$ and coordination number $N(r)$. Once *PDFgetN* is terminated, the intermediate files are deleted and only the files carrying the suffixes .sq and .gr, i.e. only the final merged $S(Q)$ and PDF that both contain the processing parameters pre-pended to the data are being saved, unless the program was started with the '-keep' option or the 'Delete files' option in the 'Options' menu was deselected. If a history (.hst) file was explicitly save using the 'Save' command under the 'File' menu this will also be kept.

A.3 History files and templates

In this appendix, the content of the history file is described. The history file contains all parameters input by the user and further information, e.g. cross-section data etc. These parameters can be modified by the user. The history file can be edited by left-clicking on *Edit all* button. Only parameters that have an influence on the processing are being described. Note, that any history file can be used as template file for *PDFgetN*. Simply remove the information that is not required for your template and copy the file in the `templates` directory of the *PDFgetN* distribution. The new template will then automatically show up in the drop-down menu of the startup screen.
The history file is reproduced block-by-block and the description of its content is given below each block. The first section contains three blocks for information related to the sample run, vanadium run and container run.

```
History written: Mon Feb 24 10:11:15 2003
produced by Thomas
##### Run Information     runCorrection=T
prep=gsas     machine=npd
run=npdf_00133     background=npdf_00132
smooth=2    smoothParam=32 32 0    backKillThresh=-1.0
in beam: radius=0.4763     height=5.00
temp=300     runTitle=Calibration - Ni

##### Vanadium     runCorrection=T
run=npdf_00130     background=npdf_00132
smooth=2    smoothParam=32 32 0    vanKillThresh=4.0    vBackKillThresh=-1.0
in beam: radius=0.4763     height=5.00

##### Container     runCorrection=T
run=npdf_00131     background=npdf_00132
smooth=2    smoothParam=32 32 0    cBackKillThresh=-1.0
wallThick=0.023     atomDensity=0.072110
atomic information: scattCS=5.100     absorpCS=5.080
```

Parameter	**Description**
runCorrection	T=Sample correction will be calculated, F=No corrections
run	Sample data file(s)
background	Instrument background for sample
smooth	0=No smoothing of background
	1=Background will be smoothed using overlapping second-order polynomials
	2=Background will be smoothed using the Savitzky-Golay filter
smoothParam	*smooth=1*: DQFIT, FITREG, DQPEAK
	DQFIT=width of fitting intervals in units of $\Delta Q/Q$

Parameter	Description
	FITREG=ratio of extrapolated points to fit points, i.e. if 1000 points are fit and FITREG=0.05 then the polynomial from the fit will be used to smooth the central 50 points of the fit region
	DQPEAK=width of vanadium Bragg peaks in units of $\Delta Q/Q$
	smooth=2: LEFT, RIGHT, POLYNOM
	LEFT=number of points left of a data point for averaging
	RIGHT=number of points right of a data point for averaging
	POLYNOM=Order of polynomial function for averaging
radius	Inner radius of the sample container (cm)
height	Height of the sample in the beam (cm)
temp	Sample temperature (K)
runTitle	Title of this experiment
wallThick	Thickness of the wall of the sample container, usually outer radius - inner radius
atomDensity	Number density of container
scattCS	Scattering cross section of container, default is for vanadium
absorbCS	Absorption cross section of container
backKillThresh	Parameter for Bragg peak removal for background run, -1: off, 0: full smoothing and peak removal, N: points greater than $N\sigma$ are treated as peak and removed
vanKillThresh	Same as above, just for vanadium run
vBackKillThresh	Same as above, just for vanadium background
cBackKillThresh	Same as above, just for container background

Next comes a section with information about the sample, elements, composition and density.

```
##### Sample Material     numElements=1     NormLaue=0.00000
Element relAtomNum atomMass atomCoherCS atomIncoherCS atomAbsorpCS
  Ni        1.0000   58.693     13.3000        5.2000      4.49000
density=4.0     effDensity=2.55
```

Parameter	Description
numElements	Number of elements in the sample
relAtomNum	Relative number density of element
atomMass	Molar mass of element
atomCoherCS	Coherent cross section of element
atomIncoherCS	Incoherent cross section of element
atomAbsorpCS	Absorption cross section of element
density	Bulk density of sample in g/cm^3 (*not used*)
effDensity	Effective density of sample in g/cm^3

Next comes a section with detector bank information.

```
##### Banks=6  deltaQ=0.01  matchRef=1  matchScal=T  matchOffset=T
bank   angle   blendQmin   blendQmax       (0.0 means no info)
  1      90.0      1.51    33.10
  2     -90.0      1.51    33.10
  3     119.0      1.84    40.34
  4    -119.0      1.84    40.34
  5     148.0      2.05    45.00
  6    -148.0      2.05    45.00
```

Parameter	**Description**
Banks	Number of banks to process
deltaQ	Grid size ΔQ for the binned data
matchRef	Reference bank number for automatic matching of $S(Q)$, 0=no matching
matchScal	T=use scale for auto-matching, F=not
matchOffset	T=use offset for auto-matching, F=not
bank	Bank number
angle	Bank angle in units of 2θ
blendQmin	Desired Q_{min} for this bank
blendQmax	Q_{max} for this bank

The next section contains information for the different processing programs.

```
##### Program Specific Information
## Ft     calcError=1     (1 for true, 0 for false)
numRpoints=10000     maxR=100.0     numDensity=0.0   intMaxR=1.5
## Damp     Qmin=1.51     Qmax=35.00     startDampQ=35.00    QAveMin=27
dampFuncType=0     modEqn=1.0000*S(Q) +0.0000 +0.0000*Q     dampExtraToZero=0
## Blend     numBanks=6     banks=1,2,3,4,5,6
## Soqd     minProcOut=0
samPlazcek=1     vanPlazcek=1     smoothData=0     modifyData=1
## Corps     minProcOut=0     numBanksMiss=0
```

Parameter	**Description**
calcError	1=propagates error in Fourier transform 0=no error propagation in Fourier transform
numRpoints	Number of points in the PDF, determines the stepsize according to maxR/numRpoints, e.g. for maxR=20.0 and numRpoints=2000 the PDF is calculated on a grid with Δr=0.01 Å.
maxR	Maximum distance for which PDF will be calculated
numDensity	number density of the sample. If $>$0 then $\rho(r)$ and $N(r)$ are calculated

Parameter	Description
intMaxR	Value of r_{low} for calculation of G_{low}
Qmin	Q$_{min}$ for final merged $S(Q)$
Qmax	Q$_{max}$ for final merged $S(Q)$
startDampQ	Data will be damped from Q$_{damp}$ to Q$_{max}$
dampFuncType	1=$D1(Q) = sin[\frac{\pi(Q-Q_{damp})}{Q_{max}-Q_{damp}} / \frac{Q-Q_{damp}}{Q_{max}-Q_{damp}}]$ 2=$D2(Q) = D1(Q)exp\{-[\frac{2(Q-Q_{damp})}{Q_{max}-Q_{damp}}]^2\}$ 4=$D4(Q) = cos^2[\frac{\pi(Q-Q_{damp})}{2(Q_{max}-Q_{damp})}]$ any other value = no damping
QAveMin	Start value Q for calculating average $< S(Q) >$
modEqn	$S(Q)$ will be modified according to $S(Q) = aS(Q) + b + cQ$ before Fourier transforming
dampExtraToZero	Controls extrapolation of $S(Q)$ to 0. If this value is smaller than 0.0 then extrapolate $S(Q)$ to $S(Q=0) = \frac{\langle\langle b\rangle\rangle^2 - \langle\langle b^2\rangle\rangle}{\langle\langle b\rangle\rangle^2}$ else $S(Q < Q_{min}) = 0.0$
numBanks	Number of banks to be included in the final $S(Q)$
banks	Bank numbers of the banks to be included
minProcOut	0=Minimum output to Log file, 1=more, 2=even more
samPlaczek	1=Calculates Placzek correction for sample, 0=not
vanPlaczek	1=Calculates Placzek correction for vanadium, 0=not
smoothData	Additional smoothing of $S(Q)$ (not used)
modifyData	1=corrections will be applied 0=no corrections will be applied

The last section of the history file depends on the data file format and the preprocessing program required. The sections for the different formats are discussed below.

GSAS

This section of the history (or template) file is specific to using the GSAS file format. The corresponding template distributed with *PDFgetN* is called `GSAS.temp`. From the drop down menu on the startup screen select **GSAS** when working with this file format.

```
##### prepgsas     prepOutput=1    numBanksMiss=0     fileExt=gsa
instParamFile=
numBanksAdd=0
numBanksMult=0
```

Parameter	Description
prepOutput	0=Minimal output to log file, 1=more, 2=even more

Parameter	Description
numBanksMiss	Number of banks to miss, if > 0 than bank numbers to be missed have to be supplied
fileExt	File extension for data files
instParamFile	Name if the GSAS instrument parameter file
numBanksAdd	Number of banks with additive corrections, if > 0 bank number and additive constants for data and background need to be specified below
numBanksMul	Number of banks with multiplicative corrections, if > 0 bank number and multiplicative constants for data and background need to be specified below

IPNS

This section of the history (or template) file is specific to using the IPNS file format. Note that the IPNS run files need to be converted to ASCII using `sepdbtoa`. The corresponding template distributed with *PDFgetN* is called `IPNS_Ascii.temp`. From the drop down menu on the startup screen select **IPNS_Ascii** when working with this file format.

```
##### prepsepd     prepOutput=1     numBanksMiss=0     fileExt=asc
firstMonitor=5     secondMonitor=7
neutFunc=1     neutFrac=0.04
neutMinLambda=0.25     neutMaxLambda=5.27
neutStartLambda=0.0     neutEndLambda=0.0
numCal=6
Angle    Difc        DifA       Tzero
139.69   7356.98     1.45      -1.84
150.0    7570.45     1.40      -2.83
90.0     5578.08     3.91       1.75
60.0     3904.24    10.77      13.04
30.0     2025.77     0.00       0.00
14.62    997.74      0.00       0.00
numBanksAdd=0
numBanksMult=0
```

Parameter	Description
numBanksMiss	number of banks to miss, if > 0 than bank numbers to be missed have to be supplied
prepsOutput	0=Minimal output to log file, 1=more, 2=even more
fileExt	File extension for data files
firstMonitor	Bank number for the monitor. **Make sure that this is correct**
secondMonitor	Bank number for the transmission monitor
neutFunc	Function for delayed neutron correction 0=no correction 1=constant fraction, needs next three parameters

Parameter	Description
	2=sawtooth correction (appropriate if delayed neutron chopper was used), needs next five parameters
neutFrac	Delayed neutron fraction
neutMinLambda	Minimum wavelength in Åfor delayed neutron correction
neutMaxLambda	Maximum wavelength in Åfor delayed neutron correction
neutStartLambda	Beginning of sawtooth cut in delayed neutron background, neutStartLambda $<$ neutMinLambda
neutEndLambda	End of sawtooth cut in delayed neutron background neutEndLambda $>$ neutMaxLambda
numCal	Number of entries for diffractometer constants to be supplied in the table following numCal. The default is for the joint Rietveld/PDF histogram. **Adjust for your histogram if necessary**
Angle	Bank angle in units of 2θ
DifC, DifA, Tzero	Diffractometer constants
numBanksAdd	Number of banks with additive correction
numBanksMult	Number of banks with multiplicative corrections

ISIS

This section of the history (or template) file is specific to using the ISIS file formats. Note that there are two different formats: *Ariel* and *Norm*. Also note that *Norm* files need to be converted to ASCII first. The corresponding templates distributed with *PDFgetN* are called `ISIS_norm.temp` and `ISIS_ariel.temp` respectively. From the drop down menu on the startup screen select **ISIS_norm** or **ISIS_ariel** when working with these file formats.

```
##### prepnorm     prepOutput=1     numBanksMiss=0     fileExt=asc
numBanksAdd=0
numBanksMult=0
```

Parameter	Description
prepOutput	0=Minimal output to log file, 1=more, 2=even more
numBanksMiss	Number of banks to miss, if >0 than bank numbers to be missed have to be supplied
fileExt	File extension for data files
numBanksAdd	Number of banks with additive corrections, if >0 bank number and additive constants for data and background need to be specified below
numBanksMul	Number of banks with multiplicative corrections, if >0 bank number and multiplicative constants for data and background need to be specified below

Appendix B

Installation

B.1 Windows

The Windows version of *PDFgetN* is distributed as a self-extracting installer. Simply download the file `PDFgetN-1.6.x-win32.exe` (or the most current version) and run it by double clicking on the downloaded file. This will start the installation process and the dialog shown in Figure B.1 will appear. Simply follow the instructions on the screen and that is all, you are ready to use *PDFgetN*. The installer also contains a version of the program *KUPLOT* required for the plotting functions of *PDFgetN*.

B.2 Linux/Unix

Unfortunately the installation of *PDFgetN* is not as simple, since the program needs to be compiled from the source and several other programs need to be installed. For a list of required programs and where to get them, refer to Table B.2. First download the archive `PDFgetN-1.6.x.tar.gz` or the most current version. Next you need to unpack the archive into a temporary directory. This is done using the commands:

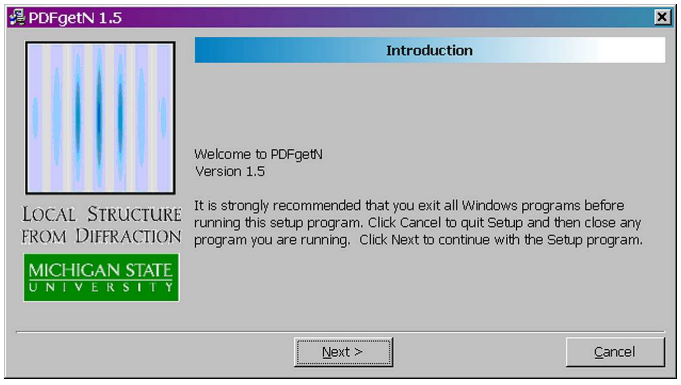

Figure B.1: Windows installer for *PDFgetN*.

Program	Description	Version
gcc/g77	C and FORTRAN compiler `http://www.gnu.org/software/gcc/gcc.html`	2.95
Perl	Perl scripting language `http://www.perl.com/pub/a/language/info/software.html`	5.x
PerlTk	Tk (GUI) package for Perl `http://www.perl.com/CPAN-local/modules/by-module/Tk/`	800.012
KUPLOT	KUPLOT plotting program `http://www.totalscattering.org/programs/discus/`	4.1⋆
OpenGenie	Only required for converting ISIS NORM files `http://www.isis.rl.ac.uk/openGenie/`	2.1⋆

Table B.1: List of programs required to install *PDFgetN* under Unix/Linux. Versions marked with ⋆ indicate optional program packages.

```
gzip -d PDFgetN-1.6.x.tar.gz
tar -xvof PDFgetN-1.6.x.tar
```

This will create a directory `PDFgetN-1.6.x`. Change to that directory using the command `cd PDFgetN-1.6.x`. Next start the installation by typing

```
perl install.pm
```

If you do not have `perl` installed, you will get and error message. Install `perl` and `perlTk` first and restart the installation. The installer will ask for an installation directory, make sure you actually have write permission in that directory. Note, that it is generally not required to log in as `root` to install *PDFgetN*. In case the installer does not detect the correct version of the required additional software, an error message will appear. Refer to Table B.2 for information how to obtain the missing programs. Note, that *PDFgetN* can be installed without *KUPLOT*, however, no plotting will then be possible. Once the installation is finished successfully, make sure the installation location is in your path, so you can start the program simply by typing `PDFgetN`.

Bibliography

S.J.L. Billinge; T. Egami. Short-Range Atomic Structure of $Nd_{2-x}Ce_xCuO_{4-y}$ Determined by Real-Space Refinement of Neutron-Powder-Diffraction Data. Phys. Rev. B **47** (1993) 14386–14406.

T. Egami. PDF Analysis applied to Crystalline Materials. In S. J. L. Billinge; M. F. Thorpe, editors, *Local Structure from Diffraction* page 1 New York 1998. Plenum.

T. Egami; S. J. L. Billinge. *Underneath the Bragg-Peaks: Structural Analysis of Complex Materials* . Elsevier Science B.V. Amsterdam 2003.

F. Frey. Diffuse Scattering from Disordered Crystals . Acta Cryst. B **51** (1995) 592–602.

A. C. Larson; R. B. Von Dreele. General Structure Analysis System (GSAS). Report LAUR 86-748 Los Alamos National Laboratory 2000.

R. L. McGreevy; L. Pusztai. Reverse Monte Carlo Simulation: a New Technique for the Determination of Disordered Structures . Mol. Simul. **1** (1988) 359–367.

V.M. Nield; D.A. Keen; R.L. McGreevy. The Interpretation of Single Crystal Diffuse Scattering using Reverse Monte Carlo Modelling . Acta Cryst. A **51** (1995) 763–771.

P. F. Peterson; D. Mikkelson; R. Mikkelson; T. Worlton; A. Chatterjee; D. Chen; J. Hammonds; C.-K. Loong J. Richardson. Integrated Spectral Analysis Workbench software (ISAW) 2003a. http://www.pns.anl.gov/computing/isaw/.

P. F. Peterson; Th. Proffen; I.-K. Jeong; S. J. L. Billinge; K.-S. Choi; M. G. Kanatzidis; P. G. Radaelli. Local Atomic Strain in $ZnSe_{1-x}Te_x$ from High Real Space Resolution Neutron Pair Distribution Function Measurements. Phys. Rev. B **63** (2001) 165211.

P. F. Peterson; E. S. Božin; Th. Proffen; S. J. L. Billinge. Improved Measures of Atomic Pair Distribution Function Quality. J. Appl. Cryst. **36** (2003) 53–64.

D. L. Price. GLASS package. internal report 14 IPNS undated.

Th. Proffen; S. J. L. Billinge. *PDFFIT*, a Program for Full Profile Structural Refinement of the Atomic Pair Distribution Function. J. Appl. Cryst. **32** (1999) 572–575.

Th. Proffen; S. J. L. Billinge; T. Egami; D. Louca. Structural analysis of complex materials using the atomic pair distribution function - a practical guide. Z. Krist. **218** (2003) 132–143.

Th. Proffen; T. Egami; S. J. L. Billinge; A. K. Cheetham; D. Louca; J. B. Parise. Building a high resolution total scattering powder diffractometer - Upgrade of NPD at MLNSC. Appl. Phys. A **74** (2002) S163–S165.

Th. Proffen; R.B. Neder. DISCUS, a Program for Diffuse Scattering and Defect Structure Simulations . J. Appl. Cryst. **30** (1997) 171.

Th. Proffen; T.R. Welberry. Analysis of Diffuse Scattering *via* Reverse Monte Carlo Technique: a Systematic Investigation . Acta Cryst. A **53** (1997) 202–216.

Th. Proffen; T.R. Welberry. Analysis of Diffuse Scattering from Single Crystals *via* Reverse Monte Carlo Technique: II. The Defect Structure of Calcium Stabilised Zirconia . J. Appl. Cryst. **31** (1998) 318–326.

H. M. Rietveld. A Profile Refinement Method for Nuclear and Magnetic Structures. J. Appl. Crystallogr. **2** (1969) 65.

SPEC. Certified Scientific Software 2003. http://www.certif.com/.

B.H. Toby; T. Egami. Accuracy of Pair Distribution Function Analysis Applied to Crystalline and Non-Crystalline Materials. Acta Cryst. A **48** (1992) 336–346.

B.E. Warren. *X-ray Diffraction* . Dover Publications Inc. New York 1990.

T.R. Welberry; B.D. Butler. Diffuse X-ray Scattering from Disordered Crystals . Chem. Rev. **95** (1995) 2369–2403.

T.R. Welberry; Th. Proffen. Analysis of Diffuse Scattering from Single Crystals *via* Reverse Monte Carlo: I. Comparison with Direct Monte Carlo . J. Appl. Cryst. **31** (1998) 309–317.